

# Sparse preconditioning for model predictive control

Andrew Knyazev<sup>1</sup> and Alexander Malyshev<sup>2</sup>

**Abstract**—We propose fast  $O(N)$  preconditioning, where  $N$  is the number of gridpoints on the prediction horizon, for iterative solution of (non)-linear systems appearing in model predictive control methods such as forward-difference Newton-Krylov methods. The Continuation/GMRES method for nonlinear model predictive control, suggested by T. Ohtsuka in 2004, is a specific application of the Newton-Krylov method, which uses the GMRES iterative algorithm to solve a forward difference approximation of the optimality equations on every time step.

## I. INTRODUCTION

The paper deals with novel sparse preconditioning for model predictive control using, as a specific example, the Continuation/GMRES method for on-line nonlinear model predictive control suggested by T. Ohtsuka in [10]. The method becomes popular in solving industrial applications; see, e.g. [3]. The paper [7] gives guidelines how to use the method in cases, when the system dynamics obeys a geometric structure, e.g. the symplectic one, or when the state lies on a smooth manifold. The structure-preserving solver may increase accuracy of the numerical solution. The paper [8] treats the problems with the particle solutions for nonlinear MPC using Continuation/GMRES.

The Continuation/GMRES method is based on Newton-type optimization schemes. The exact Newton method requires an analytic expression of a corresponding Jacobian matrix, which is rarely available in practice and is often replaced with a forward difference (FD) approximation; see, e.g., [4]. Such approximate Newton-type optimization schemes utilize the FD approximation of the original nonlinear equation at every time step. T. Ohtsuka uses the GMRES algorithm to solve a finite-difference approximation  $Ax = b$  to the optimality conditions. To cope with possible ill-conditioning of  $A$ , the authors of [15] propose a preconditioning strategy, which proved to be not very efficient.

In [5] and [6], we systematically search for better preconditioners to accelerate the GMRES and MINRES convergence in the C/GMRES method. In the present paper, we propose a sparse efficient  $O(N)$  preconditioner for this method, where  $N$  is the number of gridpoints on the prediction horizon.

Another popular approach to numerical solution of MPC problems is developed in [12], [13], [14], [16] and based on the interior-point method. The authors of [16] develop a

direct method for a linear MPC model with the  $O(N)$  arithmetic complexity. The papers [13], [14] apply the MINRES iteration with special preconditioners to similar linear MPC problems and prove the  $O(N)$  arithmetical complexity of the preconditioned iteration. In contrast to the above methods, which use the Newton or quasi-Newton approximations, the recent papers [2] and [9] investigate performance of the first-order methods and their Nesterov's acceleration.

Our proposed preconditioning technique is essentially based on two ideas. The symmetric matrix  $A$  is a Schur complement of the Hessian of the Lagrangian function, associated with the model prediction problem. Apart from a few rows and columns, the preconditioner  $M = LU$  is a sort of an incomplete LU factorization [1] of the Schur complement without these exceptional rows and columns. This results in sparse  $M$  and its factors,  $L$  and  $U$ , having only  $O(N)$  nonzero entries. On the one hand, application of this preconditioner is almost as fast as that of a diagonal preconditioner. On the other hand, our preconditioner has high quality leading to fast convergence of the iterative method, such as GMRES.

The rest of the paper is organized as follows. Section II derives the nonlinear equation (7), which solves the model prediction problem, following [5], [10]. We prove the symmetry of the Jacobian matrix for the function defining (7) in Theorem 1. Section III describes the continuation method for solving (7). Section IV formulates the preconditioned GMRES, as in [5]. Section V describes our new preconditioner. The preconditioner construction is the main result of the paper. Section VI illustrates all details of the preconditioner setup on a representative example. Section VII displays plots of numerical results.

## II. MODEL PREDICTION PROBLEM

The model predictive control (MPC) method solves a finite horizon prediction problem along a fictitious time  $\tau \in [t, t+T]$ . Our model finite horizon problem consists, following [5], [10], in choosing the control  $u(\tau)$  and parameter vector  $p$ , which minimize the performance index  $J$  as follows:

$$\min_{u,p} J,$$

where

$$J = \phi(x(\tau), p)|_{\tau=t+T} + \int_t^{t+T} L(\tau, x(\tau), u(\tau), p) d\tau$$

subject to the equation for the state dynamics

$$\frac{dx}{d\tau} = f(\tau, x(\tau), u(\tau), p), \quad (1)$$

<sup>1</sup>Andrew Knyazev is with Mitsubishi Electric Research Laboratories (MERL) 201 Broadway, 8th floor, Cambridge, MA 02139, USA knyazev@merl.com

<sup>2</sup>Alexander Malyshev is with the Mitsubishi Electric Research Laboratories (MERL) 201 Broadway, 8th floor, Cambridge, MA 02139, USA malyshev@merl.com

and the constraints

$$C(\tau, x(\tau), u(\tau), p) = 0, \quad (2)$$

$$\psi(x(\tau), p)|_{\tau=t+T} = 0. \quad (3)$$

The initial value condition  $x(\tau)|_{\tau=t}$  for (1) is the state vector  $x(t)$  of the dynamic system. The control vector  $u = u(\tau)$ , solving the problem over the horizon, is used as an input to control the system at time  $t$ . The components of the vector  $p(t)$  are parameters of the system. Equation (1) describes the system dynamic that may be nonlinear in  $x$  and  $u$ . Equations (2) and (3) give equality constraints for the state  $x$  and the control  $u$ . The horizon time length  $T$  may in principle also depend on  $t$ .

The continuous formulation of the finite horizon problem stated above is discretized on a uniform, for simplicity of presentation, time grid over the horizon  $[t, t+T]$  partitioned into  $N$  time steps of size  $\Delta\tau$ , and the time-continuous vector functions  $x(\tau)$  and  $u(\tau)$  are replaced by their indexed values  $x_i$  and  $u_i$  at the grid points  $\tau_i$ ,  $i = 0, 1, \dots, N$ . The integral of the performance cost  $J$  over the horizon is approximated by the rectangular quadrature rule. The time derivative of the state vector is approximated by the forward difference formula. The discretized optimal control problem is formulated as follows:

$$\min_{u_i, p} \left[ \phi(x_N, p) + \sum_{i=0}^{N-1} L(\tau_i, x_i, u_i, p) \Delta\tau \right],$$

subject to

$$x_{i+1} = x_i + f(\tau_i, x_i, u_i, p) \Delta\tau, \quad i = 0, 1, \dots, N-1, \quad (4)$$

$$C(\tau_i, x_i, u_i, p) = 0, \quad i = 0, 1, \dots, N-1, \quad (5)$$

$$\psi(x_N, p) = 0. \quad (6)$$

The necessary optimality conditions for the discretized finite horizon problem are obtained by means of the discrete Lagrangian function

$$\begin{aligned} \mathcal{L}(X, U) &= \phi(x_N, p) + \sum_{i=0}^{N-1} L(\tau_i, x_i, u_i, p) \Delta\tau \\ &+ \lambda_0^T [x(t) - x_0] \\ &+ \sum_{i=0}^{N-1} \lambda_{i+1}^T [x_i - x_{i+1} + f(\tau_i, x_i, u_i, p) \Delta\tau] \\ &+ \sum_{i=0}^{N-1} \mu_i^T C(\tau_i, x_i, u_i, p) \Delta\tau + \nu^T \psi(x_N, p), \end{aligned}$$

where  $X = [x_i \ \lambda_i]^T$ ,  $i = 0, 1, \dots, N$ , and  $U = [u_i \ \mu_i \ \nu \ p]^T$ ,  $i = 0, 1, \dots, N-1$ . Here,  $\lambda$  is the costate vector,  $\mu$  is the Lagrange multiplier vector associated with the constraint (5). The terminal constraint (6) is relaxed by the aid of the Lagrange multiplier  $\nu$ . For further convenience, we also introduce the Hamilton function

$$\begin{aligned} H(t, x, \lambda, u, \mu, p) &= L(t, x, u, p) \\ &+ \lambda^T f(t, x, u, p) + \mu^T C(t, x, u, p). \end{aligned}$$

The necessary optimality conditions are the (KKT) stationarity conditions:  $\mathcal{L}_{\lambda_i} = 0$ ,  $\mathcal{L}_{x_i} = 0$ ,  $i = 0, 1, \dots, N$ ,  $\mathcal{L}_{u_j} = 0$ ,  $\mathcal{L}_{\mu_j} = 0$ ,  $i = 0, 1, \dots, N-1$ ,  $\mathcal{L}_{\nu_k} = 0$ ,  $\mathcal{L}_{p_l} = 0$ .

The KKT conditions are reformulated in terms of a mapping  $F[U, x, t]$ , where the vector  $U$  combines the control input  $u$ , the Lagrange multiplier  $\mu$ , the Lagrange multiplier  $\nu$ , and the parameter  $p$ , all in one vector:

$$U(t) = [u_0^T, \dots, u_{N-1}^T, \mu_0^T, \dots, \mu_{N-1}^T, \nu^T, p^T]^T.$$

The vector argument  $x$  in  $F[U, x, t]$  denotes the current measured or estimated state vector, which serves as the initial vector  $x_0$  in the following procedure.

- 1) Starting from the current measured or estimated state  $x_0$ , compute  $x_i$ ,  $i = 0, 1, \dots, N-1$ , by the forward recursion

$$x_{i+1} = x_i + f(\tau_i, x_i, u_i, p) \Delta\tau.$$

Then starting from

$$\lambda_N = \frac{\partial \phi^T}{\partial x}(x_N, p) + \frac{\partial \psi^T}{\partial x}(x_N, p) \nu$$

compute the costates  $\lambda_i$ ,  $i = N, N-1, \dots, 1$ , by the backward recursion

$$\lambda_i = \lambda_{i+1} + \frac{\partial H^T}{\partial x}(\tau_i, x_i, \lambda_{i+1}, u_i, \mu_i, p) \Delta\tau.$$

- 2) Calculate  $F[U, x, t]$ , using just obtained  $x_i$  and  $\lambda_i$ , as

$$F[U, x, t] = \begin{bmatrix} \frac{\partial H^T}{\partial u}(\tau_0, x_0, \lambda_1, u_0, \mu_0, p) \Delta\tau \\ \vdots \\ \frac{\partial H^T}{\partial u}(\tau_i, x_i, \lambda_{i+1}, u_i, \mu_i, p) \Delta\tau \\ \vdots \\ \frac{\partial H^T}{\partial u}(\tau_{N-1}, x_{N-1}, \lambda_N, u_{N-1}, \mu_{N-1}, p) \Delta\tau \\ C(\tau_0, x_0, u_0, p) \Delta\tau \\ \vdots \\ C(\tau_i, x_i, u_i, p) \Delta\tau \\ \vdots \\ C(\tau_{N-1}, x_{N-1}, u_{N-1}, p) \Delta\tau \\ \psi(x_N, p) \\ \frac{\partial \phi^T}{\partial p}(x_N, p) + \frac{\partial \psi^T}{\partial p}(x_N, p) \nu \\ + \sum_{i=0}^{N-1} \frac{\partial H^T}{\partial p}(\tau_i, x_i, \lambda_{i+1}, u_i, \mu_i, p) \Delta\tau \end{bmatrix}.$$

The equation with respect to the unknown vector  $U(t)$

$$F[U(t), x(t), t] = 0 \quad (7)$$

gives the required necessary optimality conditions.

*Theorem 1:* The Jacobian matrix  $F_U[U, x, t]$  is symmetric for all  $U$ ,  $x$ , and  $t$ .

*Proof:* The equation  $\mathcal{L}_X(X, U) = 0$  is always solvable with respect to  $X$  by the forward recursion for  $x_i$  and backward recursion for  $\lambda_i$ . Let us denote its solution by

$X = g(U)$ . Then  $F[U] = \mathcal{L}_U(g(U), U)$  and  $F_U = \mathcal{L}_{UU}(g(U), U) + \mathcal{L}_{UX}(g(U), U)g_U$ . Differentiation of the identity  $\mathcal{L}_U(g(U), U) = 0$  with respect to  $U$  gives the identity  $\mathcal{L}_{UU}(g(U), U) + \mathcal{L}_{UX}(g(U), U)g_U(U) = 0$ . Differentiation of the identity  $\mathcal{L}_X(g(U), U) = 0$  with respect to  $U$  gives the identity  $\mathcal{L}_{XU}(g(U), U) + \mathcal{L}_{XX}(g(U), U)g_U(U) = 0$ . Hence  $g_U = -\mathcal{L}_{XX}^{-1}(g(U), U)\mathcal{L}_{XU}(g(U), U)$  and

$$F_U[U] = \mathcal{L}_{UU}(g(U), U) - \mathcal{L}_{UX}(g(U), U)\mathcal{L}_{XX}^{-1}(g(U), U)\mathcal{L}_{XU}(g(U), U), \quad (8)$$

which is the Schur complement of the symmetric Hessian matrix of  $\mathcal{L}$  at the point  $(X, U) = (g(U), U)$ . The Schur complement of any symmetric matrix is symmetric. ■

### III. CONTINUATION ALGORITHM

The controlled system is sampled on a uniform time grid  $t_j = j\Delta t$ ,  $j = 0, 1, \dots$ . Solution of equation (7) must be found at each time step  $t_j$  on the controller board, which is a challenging part of implementation of NMPC.

Let us denote  $x_j = x(t_j)$ ,  $U_j = U(t_j)$ , and rewrite the equation  $F[U_j, x_j, t_j] = 0$  equivalently in the form

$$F[U_j, x_j, t] - F[U_{j-1}, x_j, t_j] = b_j,$$

where

$$b_j = -F[U_{j-1}, x_j, t_j]. \quad (9)$$

Using a small  $h$ , which may be different from  $\Delta t$  and  $\Delta \tau$ , we introduce the forward difference operator

$$a_j(V) = (F[U_{j-1} + hV, x_j, t_j] - F[U_{j-1}, x_j, t_j])/h. \quad (10)$$

We note that the equation  $F[U_j, x_j, t_j] = 0$  is equivalent to the equation  $a_j(\Delta U_j/h) = b_j/h$ , where  $\Delta U_j = U_j - U_{j-1}$ .

Let us denote the  $k$ -th column of the  $m \times m$  identity matrix by  $e_k$ , where  $m$  is the dimension of the vector  $U$ , and define an  $m \times m$  matrix  $A_j$  with the columns  $A_j e_k$ ,  $k = 1, \dots, m$ , given by the formula  $A_j e_k = a_j(e_k)$ . The matrix  $A_j$  is an  $O(h)$  approximation of the Jacobian matrix  $F_U[U_{j-1}, x_j, t_j]$ . The Jacobian matrix  $F_U$  is symmetric by Theorem 1.

Suppose that an approximate solution  $U_0$  to the equation  $F[U_0, x_0, t_0] = 0$  is available. The first block entry of  $U_0$  is then taken as the control  $u_0$  at the state  $x_0$ . The next state  $x_1 = x(t_1)$  is either sensor estimated or computed by the formula  $x_1 = x_0 + f(t_0, x_0, u_0)\Delta t$ ; cf. (1).

At the time  $t_j$ ,  $j > 1$ , we have the state  $x_j$  and the vector  $U_{j-1}$  from the previous time  $t_{j-1}$ . Our goal is to solve the following equation with respect to  $V$ :

$$a_j(V) = b_j/h. \quad (11)$$

Then we set  $\Delta U_j = hV$ ,  $U_j = U_{j-1} + \Delta U_j$  and choose the first block component of  $U_j$  as the control  $u_j$ . The next system state  $x_{j+1} = x(t_{j+1})$  is either sensor estimated or computed by the formula  $x_{j+1} = x_j + f(t_j, x_j, u_j)\Delta t$ .

A direct way to solve (11) is generating the matrix  $A_j$  and then solving the system of linear equations  $A_j \Delta U_j = b_j$ ; e.g., by the Gaussian elimination.

A less expensive alternative is solving (11) by the GMRES method, where the operator  $a_j(V)$  is used without explicit construction of the matrix  $A_j$ ; cf., [4], [10].

### IV. PRECONDITIONED GMRES

We recall that, for a given system of linear equations  $Ax = b$  and initial approximation  $x_0$ , GMRES constructs orthonormal bases of the Krylov subspaces  $\mathcal{K}_n = \text{span}\{r_0, Ar_0, \dots, A^{n-1}r_0\}$ ,  $n = 1, 2, \dots$ , given by the columns of matrices  $Q_n$ , such that  $AQ_n = Q_{n+1}H_n$  with the upper Hessenberg matrices  $H_n$  and then searches for approximations to the solution  $x$  in the form  $x_n = Q_n y_n$ , where  $y_n = \arg \min \|AQ_n y_n - b\|_2$ .

The convergence of GMRES may stagnate for an ill-conditioned matrix  $A$ . The convergence can be improved by preconditioning. A matrix  $M$  that is close to the matrix  $A$  and such that computing  $M^{-1}r$  for an arbitrary vector  $r$  is relatively easy, is referred to as a preconditioner. The preconditioning for the system of linear equations  $Ax = b$  with the preconditioner  $M$  formally replaces the original system  $Ax = b$  with the equivalent preconditioned linear system  $M^{-1}Ax = M^{-1}b$ . If the condition number  $\|M^{-1}A\|\|A^{-1}M\|$  of the matrix  $M^{-1}A$  is small, convergence of iterative solvers for the preconditioned system can be fast.

A typical implementation of the preconditioned GMRES is given by Algorithm 1. GMRES without preconditioning is the same algorithm with  $z = r$ . In the pseudocode, we denote by  $H_{i_1:i_2, j_1:j_2}$  the submatrix of  $H$  with the entries  $H_{ij}$  such that  $i_1 \leq i \leq i_2$  and  $j_1 \leq j \leq j_2$ .

---

#### Algorithm 1 Preconditioned GMRES( $k_{\max}$ )

---

**Input:**  $a(v)$ ,  $b$ ,  $x_0$ ,  $k_{\max}$ ,  $M$

**Output:** Solution  $x$  of  $a(x) = b$

```

1:  $r = b - a(x_0)$ ,  $z = M^{-1}r$ ,  $\beta = \|z\|_2$ ,  $v_1 = z/\beta$ 
2: for  $k = 1, \dots, k_{\max}$  do
3:    $r = a(v_k)$ ,  $z = M^{-1}r$ 
4:    $H_{1:k, k} = [v_1, \dots, v_k]^T z$ 
5:    $z = z - [v_1, \dots, v_k]H_{1:k, k}$ 
6:    $H_{k+1, k} = \|z\|_2$ 
7:    $v_{k+1} = z/\|z\|_2$ 
8: end for
9:  $y = \arg \min_y \|H_{1:k_{\max}+1, 1:k_{\max}} y - [\beta, 0, \dots, 0]^T\|_2$ 
10:  $x = x_0 + [v_1, \dots, v_{k_{\max}}]y$ 

```

---

It is a common practice to compute the LU factorization  $M = LU$  by the Gaussian elimination and then compute the vector  $M^{-1}r$  by the rule  $M^{-1}r = U^{-1}(L^{-1}r)$ .

### V. SPARSE PRECONDITIONER

Our finite horizon model prediction problem allows us to construct sparse preconditioners  $M_j$  with a particular structure. These preconditioners are highly efficient, which is confirmed by the numerical experiments described below.

We first observe that the states  $x_i$ , computed by the forward recursion, and the costates  $\lambda_i$ , computed by the subsequent backward recursion, satisfy, in practice, the following property:  $\partial x_{i_1}/\partial u_{i_2} = O(\Delta \tau)$ ,  $\partial \lambda_{i_1}/\partial u_{i_2} = O(\Delta \tau)$ ,  $\partial x_{i_1}/\partial \mu_{i_2} = 0$  and  $\partial \lambda_{i_1}/\partial \mu_{i_2} = O(\Delta \tau)$ . It is a corollary of theorems about the derivatives of solutions of

ordinary differential equations with respect to a parameter; see, e.g., [11].

Now we assume that the predicted states  $x_i$  and costates  $\lambda_i$  are computed by the forward and backward recursions for the vector  $U_{j-1}$  at the current system state  $x_j = x(t_j)$  during computation of the right-hand side vector  $b_j$  and use the predicted  $x_i$  and  $\lambda_i$  to form the blocks  $H_{uu}$ ,  $H_{u\mu}$ ,  $H_{\mu u}$ ,  $H_{\mu\mu}$  of the symmetric matrix

$$M_j = \begin{bmatrix} H_{uu}(U_{j-1}, x_j, t_j) & H_{u\mu}(U_{j-1}, x_j, t_j) & M_{13} \\ H_{\mu u}(U_{j-1}, x_j, t_j) & H_{\mu\mu}(U_{j-1}, x_j, t_j) & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix},$$

where  $[M_{31}, M_{32}, M_{33}]$  coincides with the last  $l$  rows of  $A_j$ . The integer  $l$  denotes the sum of dimensions of  $\psi$  and  $p$ .

In the notation of Theorem 1, the above construction is explained as follows. We discard the second term in formula 8 and use the truncated expression  $F_U[U] = \mathcal{L}_{UU}(g(U), U)$  for the entries of  $M_j$  apart from the last  $l$  columns and last  $l$  rows. The last  $l$  columns are computed exactly, the last  $l$  rows equal the transposed last  $l$  columns because of the symmetry of  $M_j$ . The possibility to use the truncated expression is due to the above observation that  $\partial x_{i1}/\partial u_{i2} = O(\Delta\tau)$ ,  $\partial \lambda_{i1}/\partial u_{i2} = O(\Delta\tau)$ ,  $\partial x_{i1}/\partial \mu_{i2} = 0$ ,  $\partial \lambda_{i1}/\partial \mu_{i2} = O(\Delta\tau)$ . Moreover, the norm of  $A_j - M_j$  is of order  $O(\Delta\tau)$ .

The matrix  $M_j$  is sparse since the blocks  $H_{uu}$ ,  $H_{u\mu}$ ,  $H_{\mu u}$ ,  $H_{\mu\mu}$  are block diagonal and  $l$  is small. The particular structure of  $M_j$  is convenient for efficient LU factorization. It is possible to simultaneously permute the rows and columns of  $M_j$  and to obtain an arrow-like pattern of nonzero elements, which admits a fast LU factorization. A representative example of the sparse preconditioners  $M_j$  and their LU factorization is given in the next section.

As a result, the setup of  $M_j$ , computation of its LU factorization, and application of the preconditioner all cost  $O(N)$  floating point operations. The memory requirements are also of order  $O(N)$ .

## VI. EXAMPLE

We consider a test nonlinear problem, which describes the minimum-time motion from a state  $(x_0, y_0)$  to a state  $(x_f, y_f)$  with an inequality constrained control:

- State vector  $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$  and input control  $\vec{u} = \begin{bmatrix} u \\ u_d \end{bmatrix}$ .
- Parameter variables  $\vec{p} = [t_f]$ , where  $t_f$  denotes the length of the evaluation horizon.
- Nonlinear dynamics is governed by the system of ODE

$$\dot{\vec{x}} = f(\vec{x}, \vec{u}, \vec{p}) = \begin{bmatrix} (Ax + B) \cos u \\ (Ax + B) \sin u \end{bmatrix}.$$

- Constraints:  $C(\vec{x}, \vec{u}, \vec{p}) = [(u - c_u)^2 + u_d^2 - r_u^2] = 0$ , where  $c_u = c_0 + c_1 \sin(\omega t)$ , i.e., the control  $u$  always stays within the curvilinear band  $c_u - r_u \leq u \leq c_u + r_u$ .
- Terminal constraints:  $\psi(\vec{x}, \vec{p}) = \begin{bmatrix} x - x_f \\ y - y_f \end{bmatrix} = 0$  (the state should pass through the point  $(x_f, y_f)$  at  $t = t_f$ )
- Objective function to minimize:

$$J = \phi(\vec{x}, \vec{p}) + \int_t^{t+t_f} L(\vec{x}, \vec{u}, \vec{p}) dt,$$

where

$$\phi(\vec{x}, \vec{p}) = t_f, \quad L(\vec{x}, \vec{u}, \vec{p}) = -w_d u_d$$

(the state should arrive at  $(x_f, y_f)$  in the shortest time; the function  $L$  serves to stabilize the slack variable  $u_d$ )

- Constants:  $A = B = 1$ ,  $x_0 = y_0 = 0$ ,  $x_f = y_f = 1$ ,  $c_0 = 0.8$ ,  $c_1 = 0.3$ ,  $\omega = 20$ ,  $r_u = 0.2$ ,  $w_d = 0.005$ .

The horizon  $[t, t+t_f]$  is parameterized by the affine mapping  $\tau \rightarrow t + \tau t_f$  with  $\tau \in [0, 1]$ .

The components of the corresponding discretized problem on the horizon are given below:

- $\Delta\tau = 1/N$ ,  $\tau_i = i\Delta\tau$ ,  $c_{ui} = c_0 + c_1 \sin(\omega(t + \tau_i p))$ ;
- the participating variables are the state  $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ , the costate  $\begin{bmatrix} \lambda_{1,i} \\ \lambda_{2,i} \end{bmatrix}$ , the control  $\begin{bmatrix} u_i \\ u_{di} \end{bmatrix}$ , the Lagrange multipliers  $\mu_i$  and  $\begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix}$ , the parameter  $p$ ;
- the state is governed by the model equation

$$\begin{cases} x_{i+1} = x_i + \Delta\tau [p(Ax_i + B) \cos u_i], \\ y_{i+1} = y_i + \Delta\tau [p(Ax_i + B) \sin u_i], \end{cases}$$

where  $i = 0, 1, \dots, N-1$ ;

- the costate is determined by the backward recursion ( $\lambda_{1,N} = \nu_1$ ,  $\lambda_{2,N} = \nu_2$ )

$$\begin{cases} \lambda_{1,i} = \lambda_{1,i+1} \\ \quad + \Delta\tau [pA(\cos u_i \lambda_{1,i+1} + \sin u_i \lambda_{2,i+1})], \\ \lambda_{2,i} = \lambda_{2,i+1}, \end{cases}$$

where  $i = N-1, N-2, \dots, 0$ ;

- the equation  $F(U, x_0, y_0, t) = 0$ , where

$$U = [u_0, u_{d,0}, \dots, u_{N-1}, u_{d,N-1}, \mu_0, \dots, \mu_{N-1}, \nu_1, \nu_2, p],$$

has the following rows from the top to bottom:

$$\begin{cases} \Delta\tau [p(Ax_i + B)(-\sin u_i \lambda_{1,i+1} + \cos u_i \lambda_{2,i+1}) \\ \quad + 2(u_i - c_{ui})\mu_i] = 0 \\ \Delta\tau [2\mu_i u_{di} - w_d p] = 0 \end{cases}$$

$$\{\Delta\tau [(u_i - c_{ui})^2 + u_{di}^2 - r_u^2] = 0$$

$$\begin{cases} x_N - x_r = 0 \\ y_N - y_r = 0 \end{cases}$$

$$\begin{cases} \Delta\tau \left[ \sum_{i=0}^{N-1} (Ax_i + B)(\cos u_i \lambda_{1,i+1} + \sin u_i \lambda_{2,i+1}) \right. \\ \quad \left. - 2(u_i - c_{ui})\mu_i c_1 \cos(\omega(t + \tau_i p))\omega\tau_i \right. \\ \quad \left. - w_d u_{di} \right] + 1 = 0. \end{cases}$$

The matrices  $A_j$  have the sparsity structure as in Fig. 4. The preconditioner  $M_j$  is the symmetric matrix

$$M_j = \begin{bmatrix} M_{11} & 0 & M_{13} & M_{14} & M_{15} \\ 0 & M_{22} & M_{23} & 0 & M_{25} \\ M_{31} & M_{32} & 0 & 0 & M_{35} \\ M_{41} & 0 & 0 & 0 & M_{45} \\ M_{51} & M_{52} & M_{53} & M_{45} & M_{55} \end{bmatrix}$$

having the diagonal blocks  $M_{11}$ ,  $M_{13} = M_{31}^T$ ,  $M_{22}$ ,  $M_{23} = M_{32}^T$ . The diagonal entries of  $M_{11}$  equal  $\Delta\tau[2\mu_i - p(Ax_i + B)(\cos u_i \lambda_{1,i+1} + \sin u_i \lambda_{2,i+1})]$ . The diagonal entries of  $M_{22}$  equal  $\Delta\tau 2\mu_i$ . The diagonal entries of  $M_{13}$  equal  $\Delta\tau 2(u_i - c_{ui})$ . The diagonal entries of  $M_{23}$  equal  $\Delta\tau 2u_{di}$ . The entries of the vector  $M_{15}$  equal  $\Delta\tau(Ax_i + B)(-\sin u_i \lambda_{1,i+1} + \cos u_i \lambda_{2,i+1}) - \Delta\tau 2\mu_i c_1 \cos(\omega(t + \tau_i p))\omega\tau_i$ . The entries of the vector  $M_{25}$  equal  $-\Delta\tau w_d$ . The entries of the vector  $M_{35}$  equal  $-2\Delta\tau(u_i - c_{ui})c_1 \cos(\omega(t + \tau_i p))\omega\tau_i$ .

The blocks  $M_{14}$ ,  $M_{45}$ , and  $M_{55}$  equal to the respective blocks of  $A$  and have to be computed exactly. The sparsity pattern of  $M_j$  is displayed in Fig. 4.

To compute the LU factorization of  $M_j$  with  $O(N)$  floating point operations, we first repartition  $M_j$  as

$$M_j = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}, K_{11} = \begin{bmatrix} M_{11} & 0 & M_{13} \\ 0 & M_{22} & M_{23} \\ M_{31} & M_{32} & 0 \end{bmatrix},$$

where  $K_{11}$  is usually nonsingular. Using the representation

$$K_{11}^{-1} = \begin{bmatrix} M_{23}M_{32} & -M_{13}M_{32} & M_{13}M_{22} \\ -M_{23}M_{31} & M_{13}M_{31} & M_{11}M_{23} \\ M_{22}M_{31} & M_{11}M_{32} & -M_{11}M_{22} \end{bmatrix} \times \begin{bmatrix} D & & \\ & D & \\ & & D \end{bmatrix},$$

where  $D = (M_{11}M_{23}M_{32} + M_{13}M_{22}M_{31})^{-1}$ , we obtain the block triangular factors

$$L = \begin{bmatrix} I & 0 \\ K_{21}K_{11}^{-1} & I \end{bmatrix}, U = \begin{bmatrix} K_{11} & K_{12} \\ 0 & S_{22} \end{bmatrix},$$

where  $S_{22} = K_{22} - K_{21}K_{11}^{-1}K_{12}$ . The application of the preconditioner costs  $O(N)$  operations.

An alternative construction of the LU factorization uses a suitable simultaneous permutation of the rows and columns of  $M_j$  with the permutation indices  $1, 1 + N, 1 + 2N, \dots, i, i + N, i + 2N, \dots, N, 2N, 3N, 1 + 3N, 2 + 3N, 3 + 3N$ . The sparsity patterns of the permuted matrix and its lower triangular factor  $L$  are displayed in Fig. 5, the sparsity pattern of the upper triangular factor  $U$  is the transpose of that of the factor  $L$ .

## VII. NUMERICAL RESULTS

In our numerical experiments, carried out in MATLAB, the system of weakly nonlinear equations (11) for the test problem from Section VI is solved by the GMRES method. The error tolerance in GMRES is  $tol = 10^{-5}$ . The number of grid points on the horizon is  $N = 100$ , the sampling time of the simulation is  $\Delta t = 1/500$ , and  $h = 10^{-8}$ .

The sparse preconditioners for GMRES are constructed as in Section VI, and the LU factorization is computed as proposed in the last paragraph of Section VI.

Figure 1 shows the computed trajectory for the test example and Figure 2 shows the optimal control by the MPC approach using GMRES with preconditioning.

GMRES with preconditioning executes only 2 iterations at each step while keeping  $\|F\|_2$  close to  $10^{-4}$ . For comparison, we show the number of iterations in GMRES without preconditioning in Figure 3, which is 4-14 times larger.

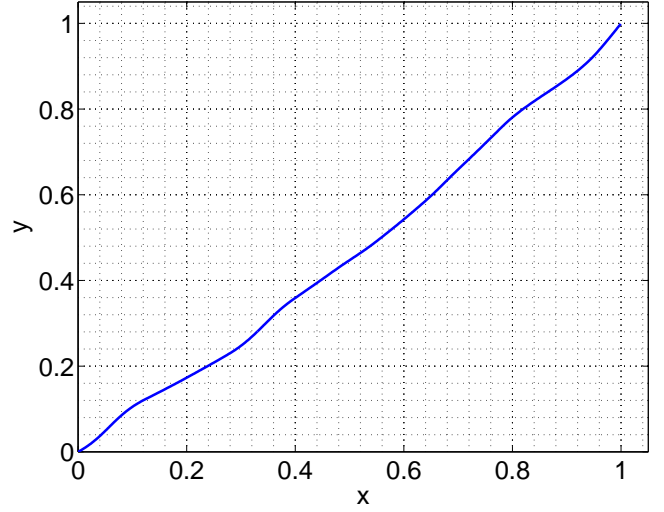


Fig. 1. Trajectory by NPMC using GMRES with preconditioning

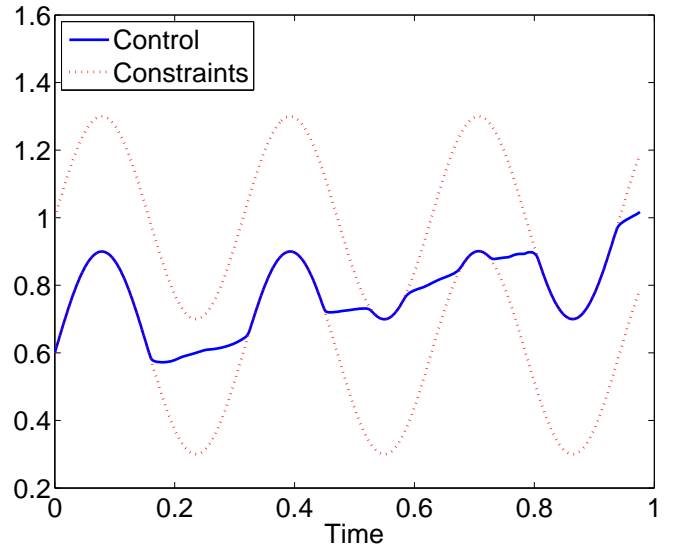


Fig. 2. NMPC control  $u$  using GMRES with preconditioning

## VIII. CONCLUSION

We propose an efficient sparse preconditioner for the Continuation/GMRES method for nonlinear MPC problems. The arithmetical cost of preconditioning is  $O(N)$ , memory storage is  $O(N)$ , where  $N$  is the number of gridpoints on the prediction horizon.

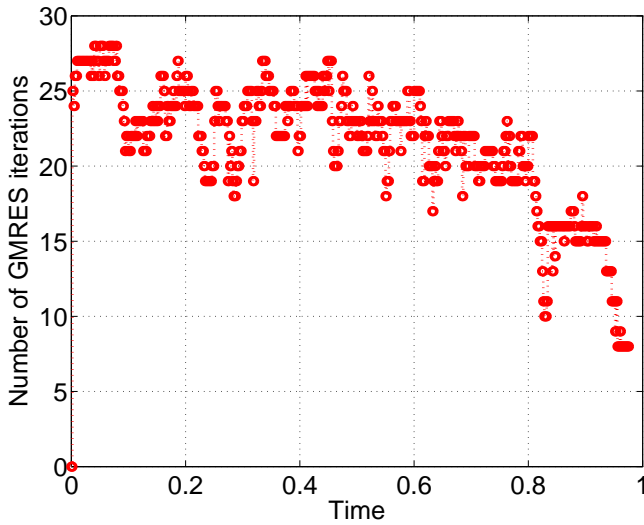


Fig. 3. The number of GMRES iterations without preconditioning,  $N = 100$ ,  $\Delta t = 1/500$ ,  $k_{\max} = 100$

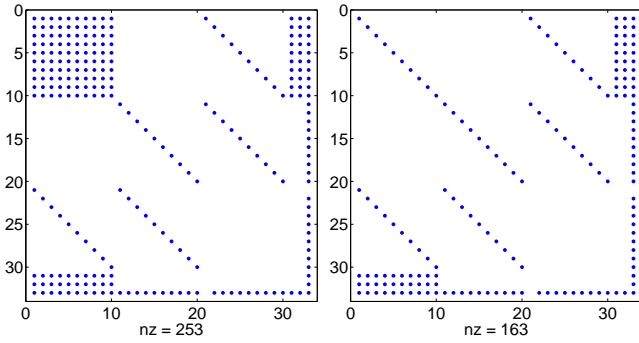


Fig. 4. Sparsity patterns of the Jacobian  $F_U$  and preconditioner  $M$

## REFERENCES

- [1] O. Axelsson, Iterative solution methods. New York, NY: Cambridge U. Press, 1994.
- [2] P. Giselsson, Optimal preconditioning and iteration complexity bounds for gradient-based optimization in model predictive control, in Proc. American Control Conf., Washington D.C., USA, June, 2013, pp. 358–364.
- [3] M. Huang, H. Nakada, K. R. Butts, and I. V. Kolmanovsky, Nonlinear Model Predictive Control of a Diesel Engine Air Path: A Comparison of Constraint Handling and Computational Strategies, in Proc. 5th IFAC Conf. Nonlinear Model Predictive Control, Seville, Spain, September 17–20, 2015, pp. 372–379.
- [4] C. T. Kelly, Iterative methods for linear and nonlinear equations. Philadelphia, PA: SIAM, 1995.
- [5] A. Knyazev, Y. Fujii, and A. Malyshev, Preconditioned Continuation Model Predictive Control, in Proc. SIAM Conf. Control Appl., Paris, France, July 8–10, 2015, pp. 101–108.
- [6] A. Knyazev and A. Malyshev, Preconditioning for Continuation Model Predictive Control, in Proc. 5th IFAC Conf. Nonlinear Model Predictive Control, Seville, Spain, September 17–20, 2015, pp. 191–196. Available at arXiv:1509.02861.
- [7] A. Knyazev and A. Malyshev, Continuation model predictive control on smooth manifolds, in Proc. 16th IFAC Workshop Control Applications of Optimization, Garmisch-Partenkirchen, Germany, October 6–9, 2015. Available at arXiv:1509.02848.
- [8] A. Knyazev and A. Malyshev, Efficient particle continuation model

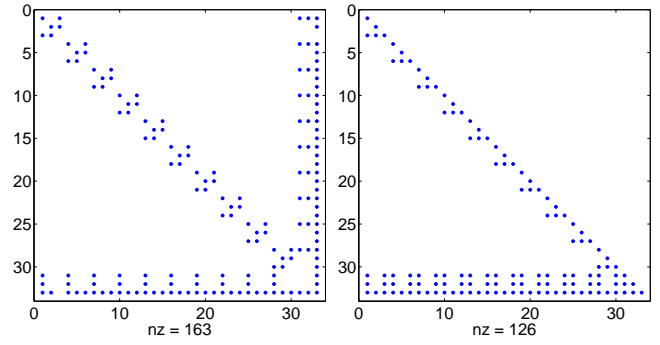


Fig. 5. Sparsity patterns of the permuted preconditioner and the lower triangular factor  $L$

- predictive control, in Proc. 16th IFAC Workshop Control Applications of Optimization, Garmisch-Partenkirchen, Germany, October 6–9, 2015. Available at arXiv:1509.02852.
- [9] D. Kouzoupis, H. J. Ferreau, H. Peyrl, and M. Diehl, First-order methods in embedded nonlinear model predictive control, in Proc. European Control Conf., Linz, Austria, 2015, pp. 2622–2627.
- [10] T. Ohtsuka, A Continuation/GMRES method for fast computation of nonlinear receding horizon control, Automatica, vol. 40, pp. 563–574, 2004.
- [11] L. S. Pontryagin, Ordinary differential equations. Reading, MA: Addison-Wesley, 1962.
- [12] C. V. Rao, S. J. Wright, and J. B. Rawlings, Application of interior-point methods to model predictive control, J. Optimiz. Theory Appl., vol. 99, pp. 723–757, 1998.
- [13] A. Shahzad, E. C. Kerrigan, and G. A. Constantinides, Preconditioners for Inexact Interior Point Methods for Predictive Control, in Proc. American Control Conf., Baltimore, MD, USA, June 30–July 2, 2010, pp. 5714–5719.
- [14] A. Shahzad, E. C. Kerrigan, and G. A. Constantinides, A Fast Well-conditioned Interior Point Method for Predictive Control, in Proc. IEEE Conf. Decision Control, Atlanta, GA, USA, December 15–17, 2010, pp. 508–513.
- [15] T. Tanida and T. Ohtsuka, Preconditioned C/GMRES algorithm for nonlinear receding horizon control of hovercrafts connected by a string, in Proc. IEEE Int. Conf. Control Appl., Taipei, Taiwan, September 2–4, 2004, pp. 1609–1614.
- [16] Y. Wang and S. Boyd, Fast model predictive control using online optimization, IEEE Trans. Control Syst. Technology, vol. 18, pp. 267–278, 2010.